A Search for Technosignatures from TRAPPIST-1, LHS 1140, and 10 Planetary Systems in the *Kepler* Field with the Green Bank Telescope at 1.15–1.73 GHz

Pavlo Pinchuk[1] , Jean-Luc Margot[1,2] , Adam H. Greenberg[1] , Thomas Ayalde[1], Chad Bloxham[1], Arjun Boddu[3],
Luis Gerardo Chinchilla-Garcia[1], Micah Cliffe[3], Sara Gallagher[1], Kira Hart[1], Brayden Hesford[3], Inbal Mizrahi[1], Ruth Pike[4],
Dominic Rodger[4], Bade Sayki[1], Una Schneck[2], Aysen Tan[3], Yinxue "Yolanda" Xiao[5], and Ryan S. Lynch[6,7]

[1] Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA; jlm@astro.ucla.edu
[2] Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA 90095, USA
[3] Department of Electrical Engineering, University of California, Los Angeles, CA 90095, USA
[4] Department of Earth Science and Engineering, Imperial College London, UK
[5] Department of Computer Science, University of California, Los Angeles, CA 90095, USA
[6] Green Bank Observatory, P.O. Box 2, Green Bank, WV 24494, USA
[7] Center for Gravitational Waves and Cosmology, Department of Physics and Astronomy, West Virginia University, White Hall, Box 6315, Morgantown,
WV 26506, USA

Abstract

As part of our ongoing search for technosignatures, we collected over three terabytes of data in 2017 May with the *L*-band receiver (1.15–1.73 GHz) of the 100 m diameter Green Bank Telescope. These observations focused primarily on planetary systems in the *Kepler* field, but also included scans of the recently discovered TRAPPIST-1 and LHS 1140 systems. We present the results of our search for narrowband signals in this data set with techniques that are generally similar to those described by Margot et al. Our improved data processing pipeline classified over 98% of the approximately six million detected signals as anthropogenic radio frequency interference (RFI). Of the remaining candidates, 30 were detected outside of densely populated frequency regions attributable to RFI. These candidates were carefully examined and determined to be of terrestrial origin. We discuss the problems associated with the common practice of ignoring frequency space around candidate detections in radio technosignature detection pipelines. These problems include inaccurate estimates of figures of merit and unreliable upper limits on the prevalence of technosignatures. We present an algorithm that mitigates these problems and improves the efficiency of the search. Specifically, our new algorithm increases the number of candidate detections by a factor of more than four compared to the results of Margot et al.

1. Introduction

The question "Are we alone in the universe?" is one of the most enduring and fundamental unanswered questions in modern science. The quest for a definitive answer to this question motivates the search for evidence of life in the solar system and beyond. Over time, two primary strategies for the search have emerged. One focuses on biosignatures, which are defined as scientific evidence of past or present life. The other focuses on technosignatures, which are defined as any scientific evidence of the existence of extinct or extant technology. Given our present knowledge of astrobiology, it is impossible to reliably predict which strategy will succeed first. We argue that the search for technosignatures offers four advantages compared to telescopic or robotic searches for biosignatures.

First, present-day human technology limits the targets for a biosignature search to planetary systems around a few dozen nearby stars, including our own Sun. In contrast, the number of available targets for a directed technosignature search is likely at least a billion times larger, given the abundance of potentially habitable exoplanets (Borucki 2016) and our ability to detect technosignatures emitted thousands of light years away (e.g., Margot et al. 2018). Second, the detection of a biosignature outside the solar system, along with its interpretation, may remain ambiguous and controversial for many years, whereas a confirmed detection of an extraterrestrial

technosignature of the type described in this work would offer a high level of certainty in interpretation. Third, the information content of some biosignatures may remain limited post-detection, while the content of an intentional extraterrestrial transmission could potentially provide an unparalleled advance in knowledge. Finally, we estimate that a substantial search for technosignatures can be accomplished with only a small (<1%) fraction of the budget currently allocated to the search for biosignatures, including missions like Mars 2020, *Europa Clipper*, and the *James Webb Space Telescope*.

In this work, we present a search for radio technosignatures using the *L*-band receiver of the 100 m diameter Green Bank Telescope (GBT). We scanned a total of 12 sources within a 2 hr observational window with the goal of detecting narrowband emissions. Narrowband (~10 Hz) signals are diagnostic of engineered emitters because the narrowest known sources of natural emission span a larger (~500 Hz) range of frequencies at the *L* band (e.g., Cohen et al. 1987). Our search builds on the legacy of technosignature searches performed in the period 1960–2010 (Tarter 2001; Tarter et al. 2010, and references therein) and complements some recent efforts (Siemion et al. 2013; Harp et al. 2016; Enriquez et al. 2017; Gray & Mooley 2017; Margot et al. 2018), but differs primarily in the choice of sample of the vast parameter space still left to search. In addition, we are generally sensitive to a wider range of signal drift rates ($\pm 8.86\,\mathrm{Hz\,s^{-1}}$) than the cited works, and to signals

Table 1
Target Host Stars Listed in Order of Observation

Host Star	Distance (ly)	HZ Category
Kepler-442	1115^{+62}_{-72}	1, 2
Kepler-440	851^{+52}_{-150}	2
Kepler-174	1210.04^{+63}_{-56}	3, 4
Kepler-62	1200	1, 2
Kepler-296	737^{+91}_{-59}	1, 2
Kepler-86	1128.5^{+44}_{-42}	4
Kepler-22	620	4
Kepler-283	1477.49^{+67}_{-74}	1, 2
Kepler-452	1787.34^{+395}_{-243}	2
Kepler-186	561^{+42}_{-33}	1, 2
TRAPPIST-1	39.5 ± 1.3	1, 2
LHS 1140	40.67 ± 1.37	1

Note. Distances in light years (ly) were obtained from the NASA Exoplanet Archive. HZ categories are described by Kane et al. (2016). Categories 1 and 2 refer to small ($R_p < 2R_E$) planets in the conservative and optimistic HZs, respectively, while categories 3 and 4 refer to planets of any radius in the conservative and optimistic HZs, respectively. TRAPPIST-1 and LHS 1140 were categorized on the basis of orbital radii from the NASA Exoplanet Archive and HZ boundaries as calculated with the algorithm of Kopparapu et al. (2013).

with lower signal-to-noise-ratio (S/N > 10) than a recent large survey (Enriquez et al. 2017; S/N > 25).

We present a substantial improvement to the signal detection algorithms used in the search pipelines of Siemion et al. (2013), Enriquez et al. (2017), and Margot et al. (2018). In these pipelines, a wide (several hundred Hertz) window surrounding the frequency of each candidate detection is removed from further consideration, whether this candidate appears in a primary (Siemion et al. 2013; Margot et al. 2018) or secondary (Enriquez et al. 2017) scan.

Consequently, other legitimate signals within that window are never analyzed, which implies that both figures of merit for the completeness of the search and upper limits on the prevalence of technosignatures reported in these works are inaccurate. By adjusting the way in which the frequency space surrounding each candidate signal is handled, we no longer need to discard legitimate signals and are able to analyze a larger fraction of the frequency space. This advance is appreciable. We estimate that this improvement alone increases our detection count by a factor of more than four as compared to the results of Margot et al. (2018). The increase in the detection count of other search pipelines after implementation of a similar improvement would likely be substantial.

Our data acquisition techniques, a brief overview of which is presented in Section 2, are generally similar to those presented by Margot et al. (2018). Section 3 explains our data analysis procedures, including major changes and improvements to our data processing pipeline. The results of our search are presented in Section 4, followed by a discussion and conclusions in Sections 5 and 6, respectively.

2. Data Acquisition

2.1. Sources

For our observations, we selected 10 sources (Table 1) from the *Kepler* catalog based on habitability criteria presented by Kane et al. (2016). These criteria take into account the size and location of exoplanets with respect to the host stars. Small planets ($R_p < 2R_E$, where R_p and R_E are the radii of the planet and Earth, respectively) in the conservative and optimistic habitable zones fall under habitable zone (HZ) categories 1 and 2, respectively. Conversely, planets with any radius in the conservative and optimistic HZs fall under HZ categories 3 and 4, respectively.

For our observations, we selected all seven host stars reported by Kane et al. (2016) with confirmed planets in HZ categories 1 and 2 (Table 1). We supplemented these with three more host stars with confirmed planets in HZ categories 3 and 4 (Table 1). In addition to the 10 sources in the *Kepler* field, we observed the recently discovered TRAPPIST-1 and LHS 1140 systems. The TRAPPIST-1 system hosts seven Earth-sized, temperate exoplanets orbiting an ultra-cool dwarf star (Gillon et al. 2017). The benign equilibrium temperatures of some of the planets in the system make the prospect of liquid water on their surfaces, and thus the possibility of life, plausible. The LHS 1140 system harbors only one known planet, which orbits its M dwarf host star within the HZ (Dittmann et al. 2017).

2.2. Observations

We conducted our observations with the 100 m diameter GBT on 2017 May 4, 15:00–17:00 UT. We recorded both linear polarizations of the *L*-band receiver using the GUPPI backend in its baseband recording mode (DuPlain et al. 2008). GUPPI was configured to channelize 800 MHz of recorded bandwidth into 256 channels of 3.125 MHz each. We verified telescope pointing accuracy with standard procedures and tested signal integrity by injecting a monochromatic tone near the receiver front end.

We observed all of our targets in pairs in order to facilitate the detection and removal of signals of terrestrial origin (Section 3.4). The sources were paired by approximately minimizing the slew time of the telescope for the duration of the observing block and then taking consecutive pairs from the solution. Pairings were adjusted to eliminate any ambiguity in the direction of origin of detected signals. Specifically, we required angular separations larger than 1° (several times the ∼9 arcmin beamwidth of the GBT at the center frequency of our observations) between pair members. The pairs are listed consecutively in Table 1 (i.e., the first pair is Kepler-442 and Kepler-440, and so on). Each pair was observed for a total of ∼330 s using an "on"–"off"–"on"–"off" sequence, where "on" represents a scan of the first source in the pair, and "off" represents a scan of its partner.

3. Analysis

3.1. Data Preprocessing

After unpacking the data, we computed 2^{20}-point Fourier transforms of the digitized raw voltages, yielding a frequency resolution of $\Delta\nu = 2.98$ Hz. We chose this frequency resolution because it is small enough to provide unambiguous detections of technosignatures and large enough to examine Doppler frequency drift rates up to ~ 10 Hz s^{-1} (Section 3.2).

We modeled the bandpass response of GUPPI's 256 channels by fitting a 16-degree Chebyshev polynomial to the median bandpass response of all of channels within the operating range of the GBT *L*-band receiver (1.15–1.73 GHz), excluding channels that overlap the frequency range (1200–1341.2 MHz) of a notch filter designed to mitigate radio frequency interference (RFI) from aircraft radar. Our experience indicates that this
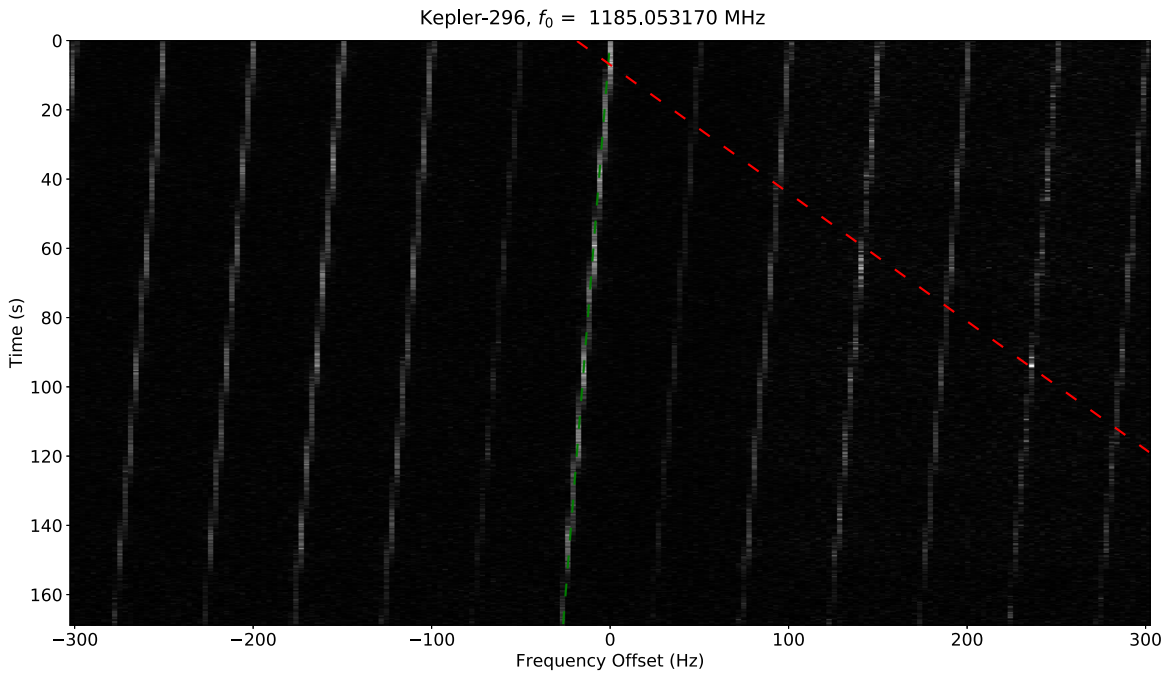
Figure 1. Time–frequency diagram of a signal that would result in multiple detections in an earlier version of our data processing pipeline. In this diagram, pixel intensity represents signal power. The green dashed line shows the correct drift rate of the signal, while the red dashed line indicates a drift rate along which signal values integrate to a summed power that exceeds our detection threshold, resulting in an additional, spurious detection. Our improved pipeline eliminates the possibility of spurious detections for signals of this nature.

process allows us to apply bandpass corrections that yield the expected flat baselines. After applying the bandpass correction to all 256 channels, we stored consecutive power spectra of length 2^{20} as rows in a time–frequency array and normalized the result to zero mean and unit variance of the noise power. We call the graphical representation of such arrays time–frequency diagrams, though they are also often referred to as spectrograms, spectral waterfalls, or waterfall plots.

3.2. Doppler De-smearing

Due to the orbital and rotational motions of both the emitter and the receiver, we expect extraterrestrial technosignatures to drift in frequency space (e.g., Siemion et al. 2013; Margot et al. 2018). A de-smearing algorithm is required to avoid spreading the power of a given signal over multiple channels. Since the Doppler drift rates due to the emitters are unknown, we examined 1023 linearly spaced drift rates in the range of $\pm 8.86\,\mathrm{Hz\,s^{-1}}$, with a step of $\Delta\dot{f} = 0.0173\,\mathrm{Hz\,s^{-1}}$. To accomplish this, we made use of a computationally advantageous Doppler de-smearing algorithm (Taylor 1974; Siemion et al. 2013), which computes an array containing de-smeared power spectra, where each de-smeared spectrum represents a time integration of the consecutive power spectra after correcting for a given Doppler drift rate. A single pass of this algorithm computes 512 power spectra for all drift rates in the range of 0–$8.86\,\mathrm{Hz\,s^{-1}}$. To obtain power spectra at negative drift rates, we applied the algorithm a second time and reversed the search direction.

In a previous analysis, Margot et al. (2018) performed a search for technosignatures on each of the resulting arrays individually. As a result, it was possible for a candidate signal to be detected twice; once with the correct drift rate and once with a spurious drift rate of the opposite sign (Figure 1). This duplication increased the false detection count and was

occasionally problematic in subsequent stages of the data processing pipeline. To avoid this problem, we concatenated the outputs of both applications of the de-smearing algorithm into a single array prior to subsequent analysis. This array contained all 1023 possible drift rates (one duplicate calculation at $0\,\mathrm{Hz\,s^{-1}}$ was removed).

3.3. Candidate Signal Detection

The output of the Doppler de-smearing algorithm contains the integrated power spectra of the scan at various drift rates, making it ideal for identifying promising candidates for extraterrestrial technosignatures (hereafter, candidate signals). We performed this search iteratively by identifying the signal with the highest S/N and storing its characteristics in a structured query language (SQL) database before moving on to the next signal. Drift rates that are similar to the drift rate with maximum S/N often yield integrated powers with large S/N values as well, and it is important to avoid redundant detections of the same signal. Siemion et al. (2013) and Margot et al. (2018) avoided the redundancy by discarding all detections within a frequency range Δf centered on the start frequency of the highest S/N signal. For the analysis at 3 Hz resolution, they defined the frequency range to be $\Delta f = 2\dot{f}_{max}\,\tau$, where $\dot{f}_{max} = 8.86\,\mathrm{Hz\,s^{-1}}$ and τ is the duration of the scan (typically around 150 s). This choice guaranteed that duplicate detections for signals with the highest detectable drift rates were not recorded. However, this procedure also removed all other valid candidates within a $\sim3000\,\mathrm{Hz}$ window of every detected candidate signal. Moreover, due to the iterative nature of the search, high S/N signals were always detected first, which prevented lower S/N signals in their vicinity from being detected. This procedure removed many legitimate candidate signals in the vicinity of higher S/N signals, ultimately leaving large regions of the spectrum unexamined. This issue results in
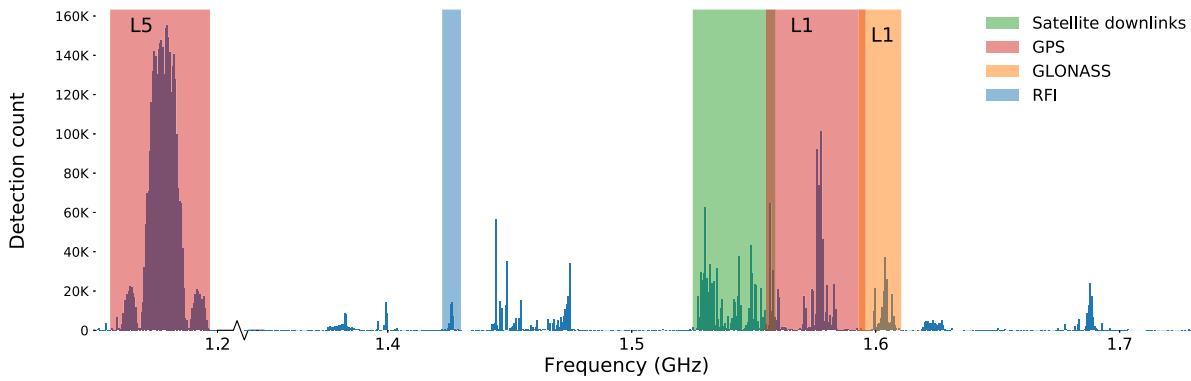
Figure 2. Candidate signal count as a function of frequency, superimposed on operating bands of known interferers. Note that the majority of detections occur in congested bands. The region marked simply as RFI contains signals from an interferer, likely air route surveillance radars, whose intermodulation products overlap the radio astronomy protected band (Margot et al. 2018). Frequencies overlapping the GBT notch filter at 1200–1341.2 MHz are excluded.

two unacceptable consequences. First, it yields incorrect calculations of figures of merit, because such calculations rely on an accurate measure of the bandwidth examined during a particular search. Second, it renders attempts to place upper limits on the abundance of technosignature sources unreliable because the pipeline eliminates the very signals it purports to detect. Problems of this nature affect the results of several studies, including those of Siemion et al. (2013), Enriquez et al. (2017), and Margot et al. (2018). To properly quantify the impact of these discarded signals on estimates of the prevalence of technosignatures, a proper injection and recovery study must be conducted. This study is beyond the scope of our work.

We designed a novel procedure to alleviate these short-comings. In order to avoid redundant detections, we simply require that no new detection contain any subset of the points in time–frequency space belonging to any other already detected signal. In other words, we require that none of our detections cross in time–frequency space. In this context, the frequency range used to discard redundant detections is no longer a constant, but rather a function of the drift rate of a new potential candidate and the bandwidth of the already detected signal. If $\dot{f}_0$ is the drift rate of a known candidate, and $\dot{f}$ is the drift rate of the potential candidate signal, then the potential candidate is marked redundant if its frequency f at the start of the scan satisfies

$$f_0 - \Delta f_b < f < f_0 + (\dot{f} - \dot{f}_0)\tau + \Delta f_b \quad \text{if} \quad \dot{f} > \dot{f}_0$$
$$f_0 + \Delta f_b > f > f_0 + (\dot{f} - \dot{f}_0)\tau - \Delta f_b \quad \text{if} \quad \dot{f} < \dot{f}_0, \quad (1)$$

where f_0 is the frequency of a known candidate signal, τ is the scan duration, and Δf_b is half of the signal bandwidth. Because we do not want detections that contain *any* part of an already detected signal, we must account for its non-zero bandwidth by extending the frequency range as in Equation (1). For implementation details of this procedure, including the estimation of bandwidth, see Appendix A.

One drawback of this method is that potential technosignatures may be discarded if they cross a stronger signal of terrestrial origin in time–frequency space. There is no reason to assume a priori that a valid candidate signal would not exhibit this behavior. However, superimposed signals are, by their very nature, difficult to detect. Other detection pipelines, including those of Siemion et al. (2013), Enriquez et al. (2017), and Margot et al. (2018), are also blind to such signals.

Table 2
Spectral Regions Exhibiting a High Density of Detections per Unit Frequency

Frequency Region (MHz)	Total Detection Count	Density (# per MHz)	Identification
1155.99–1196.91	3,579,122	87,466	GPS L5
1192.02–1212.48	96,233	4703	GLONASS L3
1422.320–1429.992	41,757	5443	Air route surveillance radar products?
1525–1559	809,877	23,820	Satellite downlinks
1554.96–1595.88	718,711	17,564	GPS L1
1592.9525–1610.485	184,207	10,507	GLONASS L1

Note. Known anthropogenic interferers are listed in the identification column.

We identified all candidate signals with S/N > 10 in channels within the operating range of the GBT *L*-band receiver (1.15–1.73 GHz), excluding channels that overlap the GBT notch filter at 1200–1341.2 MHz. A total of 5,840,149 candidate signals were detected.

3.4. Doppler and Direction-of-origin Filters

After all of the candidate signals were identified, we applied several filter procedures in order to distinguish anthropogenic signals from promising technosignature candidates. The overall purpose of the algorithms is generally similar to those described by Margot et al. (2018).

Our filter procedures (hereafter, filters) are designed to search the SQL database and flag the most promising candidates. The first filter flagged all of the candidate signals with non-zero Doppler drift rates. Signals with zero Doppler drift rates, defined here as signals that drift across less than one frequency channel over the course of a scan, are of no interest to us because the corresponding emitters will generally not be in motion with respect to the receiver, which suggests they are terrestrial in nature. The second filter flagged a signal as a promising candidate if it is persistent, i.e., it is detected in both scans of its source. This filter removes any anthropogenic signals that may have temporarily entered the beam during one of the scans. The third filter marks a signal as a technosignature candidate if its direction of origin is unique, i.e., it is not detected in the scans corresponding to other sources. This filter eliminates many anthropogenic signals that are detectable
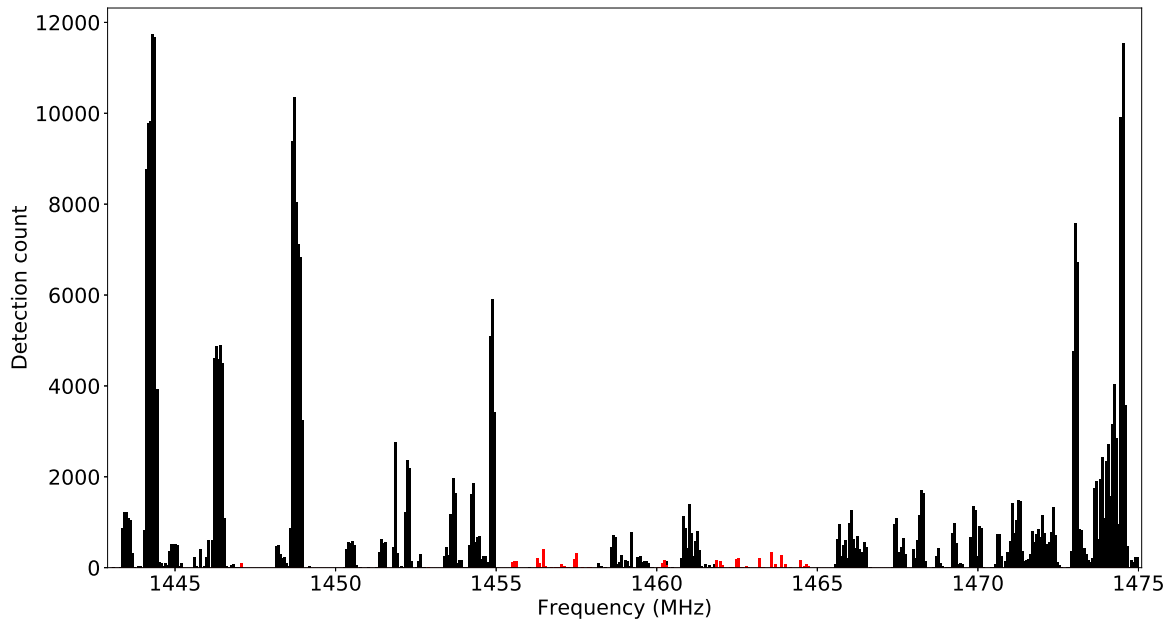
Figure 3. Example of a region of high signal density detected outside of known interferer operating bands. Signals left after the application of our signal density thresholding procedure are shown in red. Most of the remaining signals are eliminated by the Doppler and direction-of-origin filters, which are applied independently.

Table 3
Characteristics of the Top 30 Candidates

Source	Epoch (MJD)	Frequency (Hz)	Drift Rate (Hz s^{-1})	S/N	Category
Kepler-86	57877.66340278	1151559391.617775	0.0867	11.2	a
		1457412120.699883	0.4510	19.7	b
Kepler-174	57877.65087963	1501593533.158302	−0.0520	33.8	b
Kepler-186	57877.68348380	1457412961.125374	0.5551	18.6	c
		1457488715.648651	0.1908	15.6	a
		1457489272.952080	0.1735	12.9	a
		1693601790.070534	0.1041	189.0	b
Kepler-296	57877.66091435	1420354264.974594	0.2429	11.2	f
		1420476266.741753	0.2255	15.4	f
		1431607288.122177	0.2082	10.5	f
		1435340115.427971	0.1561	26.6	a
Kepler-440	57877.64343750	1457473185.658455	0.3296	26.2	b
Kepler-442	57877.64098380	1436369919.776917	0.3123	11.3	a
		1457510107.755661	0.2949	22.4	d
		1501766377.687454	−0.0520	17.8	e
LHS 1140	57877.69550926	1675002533.197403	0.01735	28.0	f
		1676205691.695213	0.01735	23.5	f
		1676211157.441139	−0.01735	24.0	f
		1677041041.851044	−0.01735	12.9	f
		1678640750.050545	0.03469	17.0	f
		1728618854.284286	−0.01735	10.0	
TRAPPIST-1	57877.69228009	1151731526.851654	0.01735	16.2	⋯
		1463927415.013313	0.03469	92.4	a
		1675910618.901253	−0.01735	14.0	f
		1676249122.619629	−0.01735	13.1	f
		1676298260.688782	−0.01735	14.6	f
		1677018946.409225	−0.01735	14.1	f
		1678110888.600349	0.03469	12.0	f
		1678460314.869881	0.01735	13.6	f
		1678716921.806335	0.01735	15.5	f

Note. Properties are listed for the first scan of the source only. For a description of the category column, see Section 4.
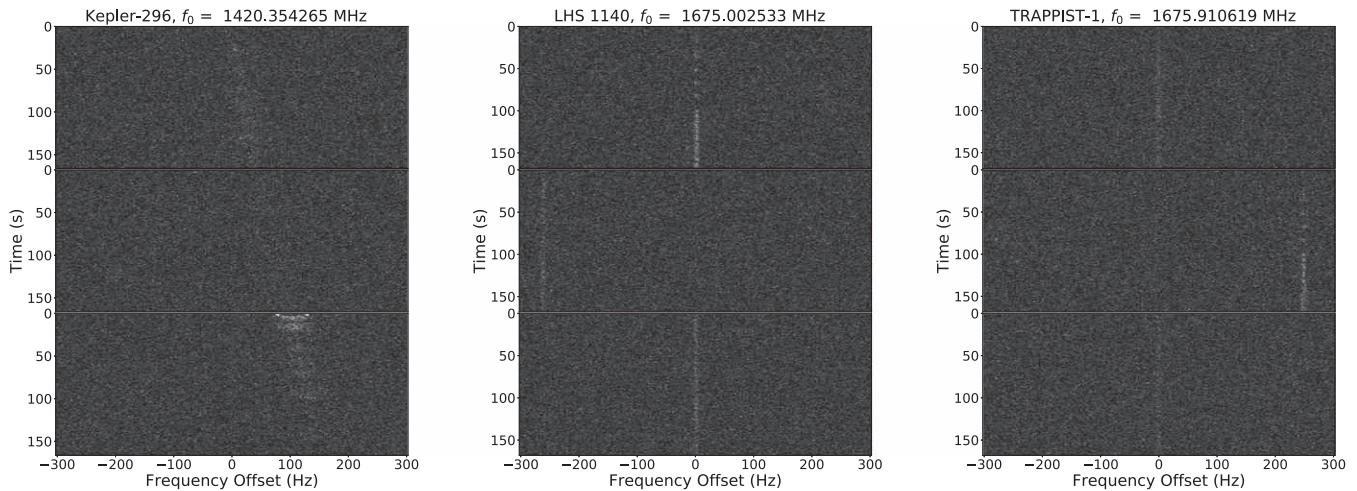
Figure 4. Time–frequency diagrams for the "on"–"off"–"on" scans for sample category "f" candidates. The signal found during the Kepler-296 scan (left) has a time–frequency structure reminiscent of the air route surveillance radar structure.

through the antenna sidelobes. For implementation details, see Appendix B.

If both a candidate signal and its corresponding signal in the other scan of the source were flagged by all three filters, then the candidate signal was marked as a high-interest signal. Of the 5,840,149 total detections, 5,743,209 (>98%) were discarded as a result of these filters.

3.5. Frequency Filters

A majority of the signals detected in our search have frequencies in operating bands of known interferers, such as global navigation satellite systems or aircraft radar. Figure 2 depicts our detection count superimposed onto frequency bands of known anthropogenic RFI. The interferer labeled simply as RFI is particularly interesting, as it overlaps the radio astronomy protected band near 1420 MHz. The time–frequency structure of this RFI is similar to that described by Siemion et al. (2013) and Margot et al. (2018), who attributed the likely origin of the RFI to intermodulation products of air route surveillance radars. Table 2 describes the properties of the regions of most prominent detected anthropogenic RFI. These regions are reminiscent of some of the permanent RFI bands of Harp et al. (2016; their Table 2). All of the candidate signals detected within these regions were removed from consideration because of their likely anthropogenic nature.

After removing candidates with frequencies in the operating bands of some well-known anthropogenic interferers (Table 2), we observed that regions of high signal density remained. An example of such a region is shown in Figure 3. Note the presence of many narrow clusters of high-signal-density regions. Although it is possible that a valid technosignature could be found within one of these regions, it would likely be difficult to detect and establish its validity given the vast amount of strong, ambient RFI. With this challenge in mind, we developed a method to discriminate signals found within densely populated frequency regions from signals found in quieter parts of the spectrum. Specifically, we measured the signal density in 1 kHz-wide frequency bins and discarded candidates within that bin if the signal density exceeded a predefined threshold value. For this work, we chose a threshold value of 1000 signals MHz^{-1}, which corresponds to a sharp drop-off in histograms of signal density. For further details, see

Appendix C. A sample result of this procedure is shown in Figure 3, where signals left after our density thresholding procedure are shown in red. More than 96% of the 581,433 signals found outside of the well-known interferer operating bands listed in Table 2 were discarded using this procedure.

4. Results

Signals that remained after the application of our Doppler, direction-of-origin, and frequency filters were marked as final technosignature candidates. Thirty such signals remained, and their properties are given in Table 3. Further examination of the final 30 candidates revealed that 13 of them are anthropogenic because they are also present in the "off" scan of the source. They were not correctly identified by our filters for a variety of reasons, which we summarize into categories below.

Category "a" refers to signals with $S/N < 10$ in the "off" scan. Because the "off" scan detections were not recorded in the database, it was not possible for our filters to flag this category of signals as RFI.

Category "b" refers to signals whose drift rates between the "on" and "off" scans differed by more than our allowed tolerance, which we set to $\pm \Delta \dot{f} = \pm 0.0173 \, \mathrm{Hz \, s^{-1}}$ (Appendix B). Because the "off" scan detections were not correctly paired to the "on" scan detections, it was not possible for our filters to flag this category of signals as RFI.

Category "c" refers to signals that cross a signal of a higher S/N in the scan of a different source. As discussed before, our signal detection methods are currently blind to such signals.

Category "d" refers to broad signals for which it is difficult to accurately determine a drift rate. Such signals are naturally difficult to pair, since the integrated power may peak at different frequencies within the bandwidth of the signal for different scans.

Category "e" refers to signals that exhibit nonlinear behavior in frequency as a function of time. Our pipeline is not currently well equipped to handle such signals because the Doppler de-smearing algorithm can only detect linear drifts in frequency versus time.

All of the categories described above pinpoint potential areas for improvement for our current pipeline.

Of the remaining 17 signals, we believe that 15 are anthropogenic. These signals appear in three sources: Kepler-296 (3),
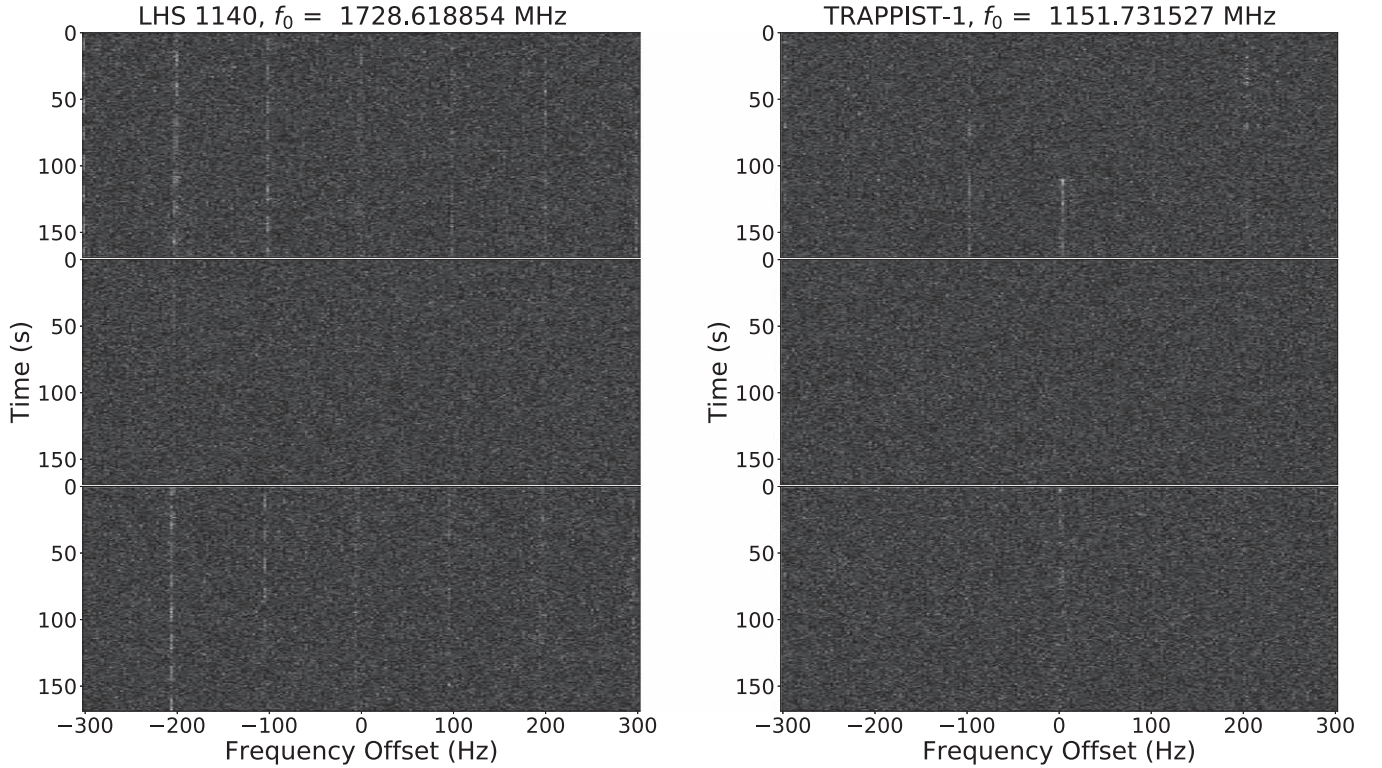
Figure 5. Time–frequency diagrams for the "on"–"off"–"on" scans for the most promising technosignature candidates. The candidate signal is at the center of each diagram. Both of the signals have a near-zero Doppler drift rate, and are in the vicinity of other signals that were discarded as anthropogenic RFI. For this reason, we cannot conclude that any extraterrestrial technosignatures were found in this search.

LHS 1140 (5), and TRAPPIST-1 (7). In each source, the signals exhibit similar modulation properties and are detected at similar frequencies but at different drift rates, implying that they are unlikely to be of extraterrestrial origin. These signals are labeled as category "f" in Table 3. The time–frequency diagrams of the "on"–"off"–"on" scans for a sample signal from each of the three sources are shown in Figure 4.

The remaining two candidates are shown in Figure 5. Note that the drift rate for both candidates is close to zero (Table 3). Additionally, both candidates are found within 100 Hz of other candidates flagged by our pipeline as RFI. As a result, we cannot conclude that these signals, nor any of the other signals found in this search, are of extraterrestrial origin.

5. Discussion

5.1. Drake Figure of Merit

The Drake figure of merit (DFM; Drake 1984) is often used to compare the parameter space examined by different searches and is given by

$$\text{DFM} = \frac{\Delta\nu_{\text{tot}}\Omega}{S_{\min}^{3/2}}, \tag{2}$$

where $\Delta\nu_{\text{tot}}$ is the total bandwidth observed, Ω is the total angular sky coverage, and $S_{\min}^{3/2}$ is the minimum flux density required for a detection. For our search, $S_{\min}^{3/2} = 9.4$ Jy and $\Omega = N \times 0.015 \deg^2$, where $N = 12$ is the number of individual sky pointings (Margot et al. 2018). To compute $\Delta\nu_{\text{tot}}$, we take the bandwidth of the GBT L-band receiver (580 MHz) and subtract the bandwidth of the GBT notch filter (141.2 MHz), the bandwidth discarded due to known interferers (Table 2;

137.167 MHz), and the total bandwidth discarded during our density thresholding procedure (Section 3.5; 37.101 MHz). Using the resulting bandwidth $\Delta\nu_{\text{tot}} = 264.532$ MHz, we find that the DFM associated with this search is 1.6×10^6. This number amounts to about 1.7% and 10% of the recent large surveys presented by Enriquez et al. (2017) and Harp et al. (2016), respectively.

5.2. Increase of Candidate Detection Efficiency

The candidate detection procedures presented by Siemion et al. (2013), Enriquez et al. (2017), and Margot et al. (2018) leave substantial regions of the spectrum unexamined, as pointed out in Section 3.3. This deficiency leads to an overestimation of the DFM associated with these searches, since this number is directly proportional to the total bandwidth examined (Equation (2)).

We can calculate the magnitude of this overestimation for the work of Margot et al. (2018) by noting that a signal with a drift rate of $\dot{f}_i$ occupies no more than $\dot{f}_i \tau$ Hz of bandwidth, where τ is the scan duration. However, a window of $f_w \approx 3000$ Hz was discarded around every detection, leaving $f_w - \dot{f}_i \tau$ Hz unexamined around every candidate signal. We can calculate the approximate fraction F by which the DFM was over-estimated using

$$F = \frac{N\Delta f_{\text{tot}}}{N\Delta f_{\text{tot}} - \sum_i (f_w - \dot{f}_i \tau)}, \tag{3}$$

where Δf_{tot} is the total bandwidth that was searched (300 MHz) and N is the total number of scans. Using the database of detections found by Margot et al. (2018), we find $F = 1.048$,
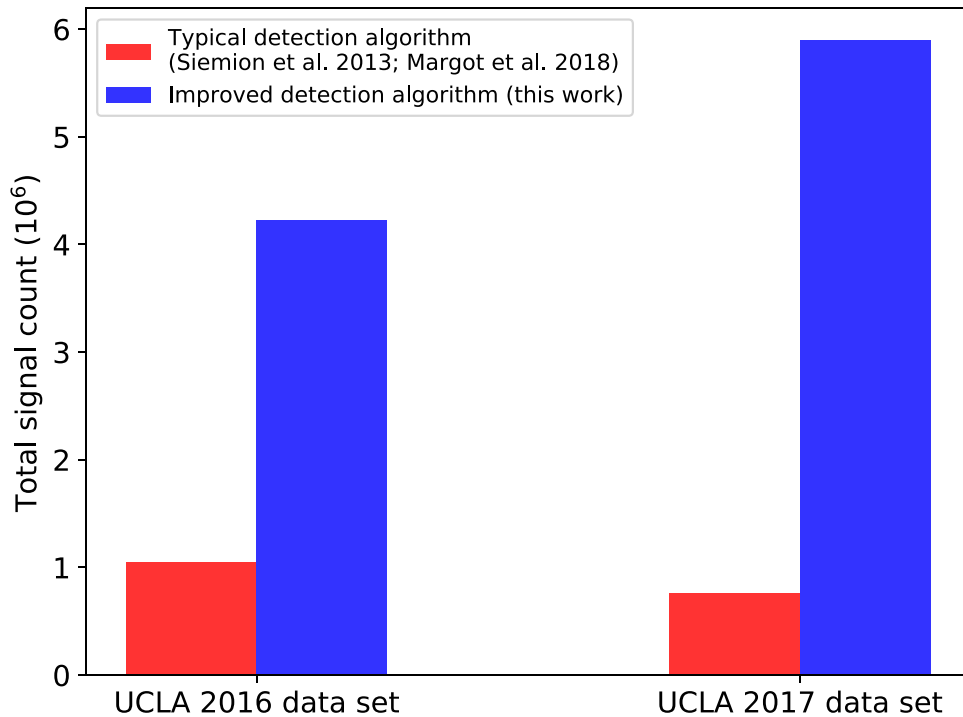
Figure 6. Detection counts obtained with the algorithm of Siemion et al. (2013), Margot et al. (2018; red), and that presented in this work (blue). The removal of many legitimate detections by typical algorithms suggests that claims of existence limits based on the results of these algorithms and others like it (e.g., Enriquez et al. 2017) are questionable.

meaning that their reported DFM was overestimated by approximately 5%. Enriquez et al. (2017) eliminate $f_w = 1200\,\text{Hz}$ of frequency space in their "on" scans centered around detections in their "off" scans. In the absence of complete information about their detections, we are unable to compute the magnitude by which their DFM was overestimated.

Our current procedure greatly improves on this situation by only requiring that detections do not overlap in time–frequency space. This improvement allows us to search the space around detected signals instead of discarding frequency windows that are 1200–3000 Hz wide. As a result, our calculation of the DFM (Section 5.1) no longer suffers from overestimation problems.

We can get a better idea of the level of improvement to our data processing pipeline by comparing the signal counts recovered by the candidate detection method used by Margot et al. (2018) and the method that is used in this work (Figure 6). Evidently, our improved detection procedure recovers more than four times the number of detections reported by Margot et al. (2018). Furthermore, 91,283 of the 97,083 signals that passed our Doppler and direction-of-origin filters (Section 3.4) were only detected as a result of this improvement. Eight of our final 30 candidates (Table 3) would not have been detected with the detection algorithms of Margot et al. (2018).

The practice of blanking frequency space around candidate detections suggests that a large number of candidates, which could include valid technosignatures, are removed from consideration when using the detection algorithms of Siemion et al. (2013), Enriquez et al. (2017), and Margot et al. (2018). This practice makes the calculation of existence upper limits unreliable, because these pipelines remove from consideration many of the signals that they are designed to detect.

We have reprocessed the 2016 data discussed by Margot et al. (2018) with our improved algorithms (Appendix D). We found no evidence of extraterrestrial technosignatures.

5.3. Existence Limits

Considering that radio technosignature detection pipelines typically eliminate substantial fractions of the spectrum (e.g., Harp et al. 2016, their Table 2; Enriquez et al. 2017, 1200–1350 MHz; Margot et al. 2018, their Table 2; this work, Table 2) and further eliminate a fraction of the signals that they are designed to detect (Siemion et al. 2013; Enriquez et al. 2017; Margot et al. 2018), it is difficult to make general and robust statements about the prevalence of narrowband emitters in the Galaxy. One such claim by Enriquez et al. (2017) has been shown to be questionable (Margot et al. 2018). Injection of artificial signals in the data would demonstrate that a fraction of detectable and legitimate signals are not identified by existing pipelines. Until this fraction is properly quantified, it is wise to refrain from making overly confident claims about the prevalence of radio emitters in the Galaxy.

5.4. Sensitivity

Margot et al. (2018) provide a detailed analysis of the sensitivity of a search performed with the 100 m GBT at a frequency resolution of $\Delta\nu = 2.98\,\text{Hz}$, sensitive to flux densities of 10 Jy. The results of that calculation (Margot et al. 2018, their Figure 5) are generally applicable here because our search parameters are identical except for a slightly better sensitivity of 9.4 Jy. We estimate that civilizations located near the closest of our observed sources (TRAPPIST-1; $\sim$40 ly) would require a transmitter with only a small fraction ($<$1%) of the effective isotropic radiated power (EIRP) of the Arecibo planetary radar to be detectable in our search. Transmitters located as far as our most-distant observed source (Kepler-452; $\sim$1800 ly) require approximately 18 times the Arecibo EIRP.

6. Conclusions

We described the results of a search for technosignatures using two hours of GBT telescope time in 2017. We identified 5,840,149 candidate signals, 98% of which were automatically eliminated by our rejection filters. Of the signals that remained, 30 were found outside of densely populated frequency regions and required further inspection. None of the remaining candidates were attributable to extraterrestrial technosignatures.

We found that quiet parts of the radio spectrum remain unexamined in the radio technosignature search pipelines of Siemion et al. (2013), Enriquez et al. (2017), and Margot et al. (2018). This problem results in inflated estimates of figures of merit and unreliable upper limits on the prevalence of technosignatures. To address this problem, we implemented a new procedure that increased the candidate detection efficiency by a factor of four or more compared to the results of Margot et al. (2018).

Our observations were designed, obtained, and analyzed by students enrolled in a UCLA course offered annually since 2016. The search for technosignatures provides a superb educational opportunity for students in astrophysics, computer science, engineering, mathematics, planetary science, and statistics. In this work, two graduate students and 15 undergraduate students at UCLA learned valuable skills related to radio astronomy, telecommunications, programming, signal processing, and statistical analysis. A course syllabus and narrative is available.[8]

Appendix A
Candidate Signal Detection and Bandwidth Estimation

To identify promising candidate signals in the drift-rate-by-frequency array output by the Doppler de-smearing algorithm, we applied an iterative procedure that seeks out candidate signals with S/N exceeding 10. We began by arranging the 1023×2^{20} array so that the drift rates decrease linearly down the rows of the array (i.e., the first row contains the integrated power with a drift rate of $8.8644 \, \mathrm{Hz \, s^{-1}}$, the second row contains the integrated power with a drift rate of $8.8471 \, \mathrm{Hz \, s^{-1}}$, and so on). We then searched for the highest S/N signal in this array, and noted its drift rate, $\dot{f}_0$, and frequency, f_0, at the start of the scan.

In order to remove redundant detections, we used Equation (1), which requires an estimate of the bandwidth of the detected signal. We measured the bandwidth of candidate signals with the Python SciPy routine `peak_widths`. These routines measure the width from the central peak of the signal to the first point on either side above a predefined threshold value, using linear interpolation when necessary. We initially measured the bandwidth as the FWHM of each signal. However, due to the peculiar nature of many of the detected signals this threshold value proved to be ineffective. We instead set the bandwidth measurement threshold at five times the standard deviation of the noise, which corresponds to half of our signal detection threshold (Figure 7).

By default, we used a search window of 200 frequency bins ($\sim$600 Hz) on either side of each signal to measure the bandwidth. For narrowband signals, this range is more than enough to ensure that the entirety of a signal fits within the search window. If a detected signal is close to the edge of the channel and the required frequency window is not available on one side of the signal, we utilize the entire available frequency range on that side of the signal, and the full 600 Hz on the other. The SciPy routines conveniently provide the frequency coordinates of the left and right intersection points at the specified threshold, which we used to calculate the width of the signal to the left (Δf_{b_l}) and right (Δf_{b_r}) of the center frequency.

Signals with large bandwidths required additional care. If the median integrated power of the signal within the initial 1200 Hz window exceeded 5σ, we labeled the signal as broadband and increased the search window to 200,000 frequency bins, or $\sim$600 kHz, on either side of the main signal. We applied a Savitzky–Golay filter (Savitzky & Golay 1964) with a window of 1001 frequency bins ($\sim$3000 Hz) and a polynomial order of 3 to the integrated spectra within the search window. This filter reduces noise by fitting a polynomial to all the points within the specified window and replacing the central point with the corresponding fit value. This filter was chosen for its computational advantages and simplicity. We then applied the bandwidth estimation procedure described above to the filtered points and stored the result in our database.

We used the bandwidth measurements in Equation (1), as follows:

$$f_0 - \Delta f_{b_l} < f < f_0 + (\dot{f} - \dot{f}_0)\tau + \Delta f_{b_r} \quad \text{if} \quad \dot{f} > \dot{f}_0$$
$$f_0 + \Delta f_{b_r} > f > f_0 + (\dot{f} - \dot{f}_0)\tau - \Delta f_{b_l} \quad \text{if} \quad \dot{f} < \dot{f}_0. \quad (4)$$

However, we chose to establish a minimum bandwidth to account for uncertainty in the determination of the drift rate of a signal (for example, if a signal's drift rate is time variable or is not a perfect multiple of the drift rate step $\Delta\dot{f}$). Specifically, we redefined Δf_{b_l} as $\min(\Delta f_{b_l}, 3\Delta\nu)$ and Δf_{b_r} as $\min(\Delta f_{b_r}, 3\Delta\nu)$, which amounts to a minimum of $\sim$10 Hz on either side of each signal for a minimum of 20 Hz per signal.

We removed redundant detections with frequencies given by Equation (4) by applying a mask to the drift-rate-by-frequency array. This mask is defined by drawing two lines through the array. The first is a vertical line at the frequency of the detected candidate signal. The second is a line with a slope of $-\Delta\dot{f}/\Delta\nu$ crossing through the frequency and drift rate point corresponding to the detected signal (Figure 8). The boundaries defined by these two lines comprise the mask for zero bandwidth signals. In order to account for the finite bandwidth of a signal, the boundaries are shifted to the left and right so
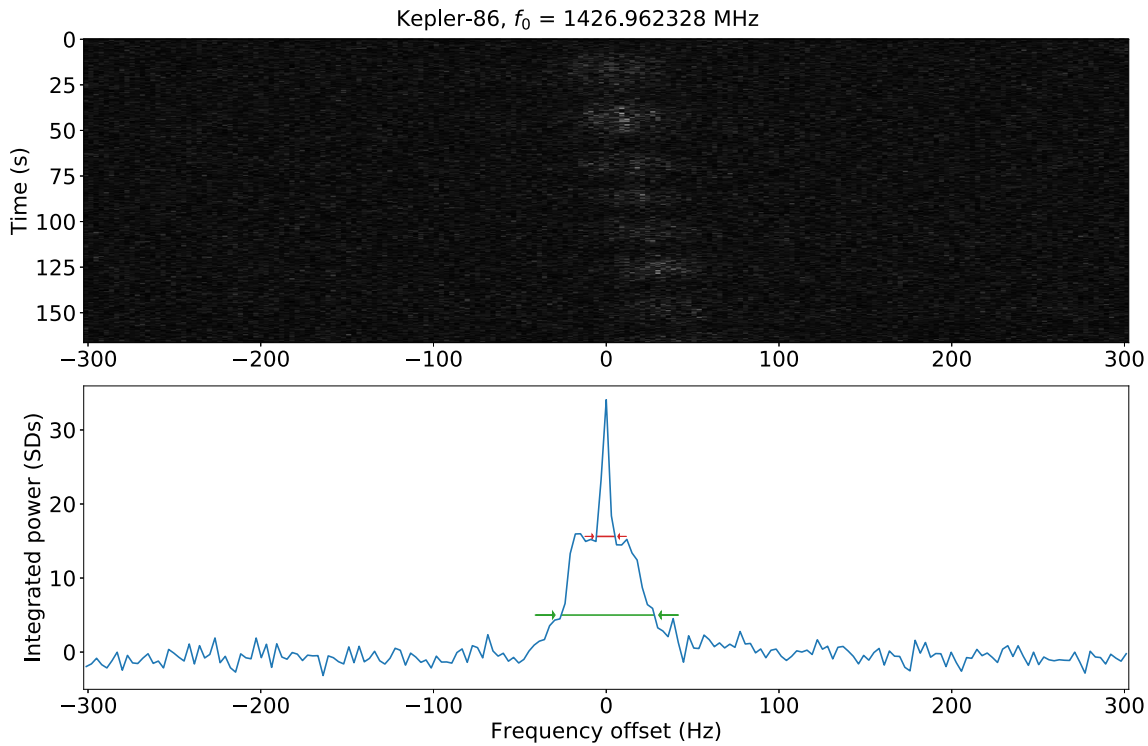
Figure 7. Sample result of the bandwidth estimation procedure. Top: time–frequency diagram of signal. Bottom: the power of the signal integrated with the best-fit drift rate of 0.2429 Hz s^{-1}, in units of standard deviations of the noise (σ). The estimated bandwidth measured at FWHM is shown by the red arrows and proved to be ineffective for the purpose of avoiding duplicate detections (e.g., signal outside this bandwidth exceeds our detection threshold of 10σ). The green arrows show the bandwidth measured at 5σ, which yields a more robust estimate of the bandwidth, which we used in our implementation.
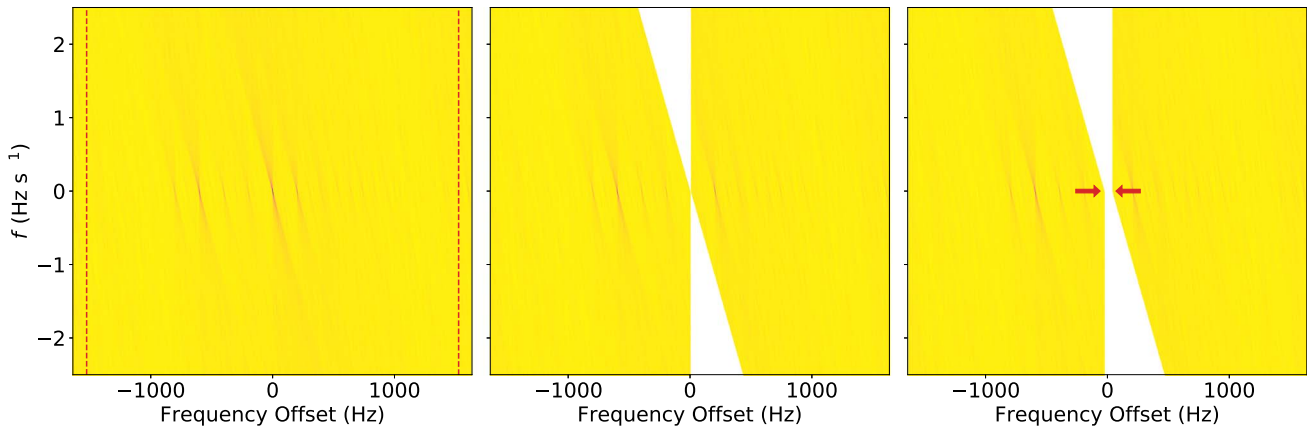

Figure 8. Left: a portion of the original array output by the Doppler de-smearing algorithm for a sample scan (LHS 1140). The plot is centered on $f = 1546.879303$ MHz. The intensity of the plot represents the integrated power at a given drift rate and frequency. Signals of interest are represented as local maxima. The minimum integrated power in this array exceeds our 10 S/N detection threshold, therefore all of the signals shown are detectable by our pipeline. The signal with maximum S/N in this portion of the array is located at the center frequency. The area between the dashed vertical lines represents the slice of the array that would have been discarded after a signal detection with the approach of Siemion et al. (2013) and Margot et al. (2018), i.e., only one of the signals would have been reported. Middle: same array after application of our new masking procedure, assuming that the measured bandwidth of the signal is 0. With our new procedure, valid signals in the vicinity of the strongest signal are not discarded. Right: same array after application of our new masking procedure, assuming that the measured bandwidth of the signal is 60 Hz. This number was chosen for visualization purposes and does not represent the true measured bandwidth ($\sim$3 Hz) of the center peak.

that the total width of the mask at the detection point matches the measured bandwidth of the signal (signified by the red arrows in Figure 8).

Appendix B
Doppler and Direction-of-origin Filters

To distinguish anthropogenic signals from potential extraterrestrial technosignatures, we invoked several filter procedures within our database to flag promising technosignature candidates. The first filter flags candidate signals with non-zero Doppler drift rates.

The direction-of-origin filters require signals from different scans to be compared and possibly paired. We pair two signals if they have similar drift rates and compatible frequencies, i.e., the frequencies at the beginning of each scan are within some tolerance of a linear relationship with a slope equal or nearly equal to the drift rate. We quantify these tests as follows.

Consider a signal with start time t_0, start frequency f_0, and drift rate $\dot{f}_0$, and another signal from a different scan with corresponding values t, f, and $\dot{f}$. We define $\Delta t = t - t_0$ and require that f falls in the interval $[f_-, f_+]$ for pairing, where

$$f_\pm = (\dot{f}_0 \pm 2\Delta\dot{f})\Delta t \pm \Delta\nu. \qquad (5)$$

In this work, the values $\Delta\dot{f}$ and $\Delta\nu$ are given by $0.0173\,\mathrm{Hz\,s^{-1}}$ and $2.98\,\mathrm{Hz}$, respectively. To account for uncertainty in the drift rate determination, we allow for a drift rate difference of $\Delta\dot{f}$. We thus query the database for all candidate signals with a frequency at the start of the scan in the range of $[f_-, f_+]$ with a drift rate of $\dot{f}_0 - \Delta\dot{f} \leqslant \dot{f} \leqslant \dot{f}_0 + \Delta\dot{f}$ $\mathrm{Hz\,s^{-1}}$. Two signals are considered paired if the following condition holds:

$$\min_i |(f - f_0) + \dot{f}_i \Delta t| \leqslant \Delta\nu, \qquad (6)$$

where $\dot{f}_i \in \{\dot{f}_0, \dot{f}_0 \pm \Delta\dot{f}, \dot{f}, \dot{f} \pm \Delta\dot{f}\}$.

To determine whether a signal is persistent, i.e., whether it is detected in both scans of its source, we apply the above procedure to each candidate and all candidates detected in the second scan of the same source. If a match is found, both signals are flagged. To determine whether a signal's direction of origin is unique, i.e., to ensure that it is not detected in the scans corresponding to other sources, we apply the above procedure to each candidate and all candidates detected in either scan of all other sources. All matches are discarded from consideration. The candidate signals that remain are flagged as having a single direction of origin.

Appendix C
Signal Density Thresholding

In order to remove candidate signals that were likely to be anthropogenic RFI, we developed a procedure to filter signals based on the density of nearby detections. We began by dividing the 1.15–1.73 GHz range into 1 kHz-wide frequency bins. For each bin, we measured the signal density by counting the number of detections within a window centered on the bin. For these calculations, we excluded the regions listed in Table 2. We tested four different window sizes: 1, 10, 100 kHz, and 1 MHz. For each window size, we plotted a histogram of signal counts. We found that a 1 MHz window resulted in a distinctive transition between small and large signal densities at approximately 1000 signals $\mathrm{MHz^{-1}}$ (Figure 9), and we used this threshold and window size to filter out regions of high signal density.
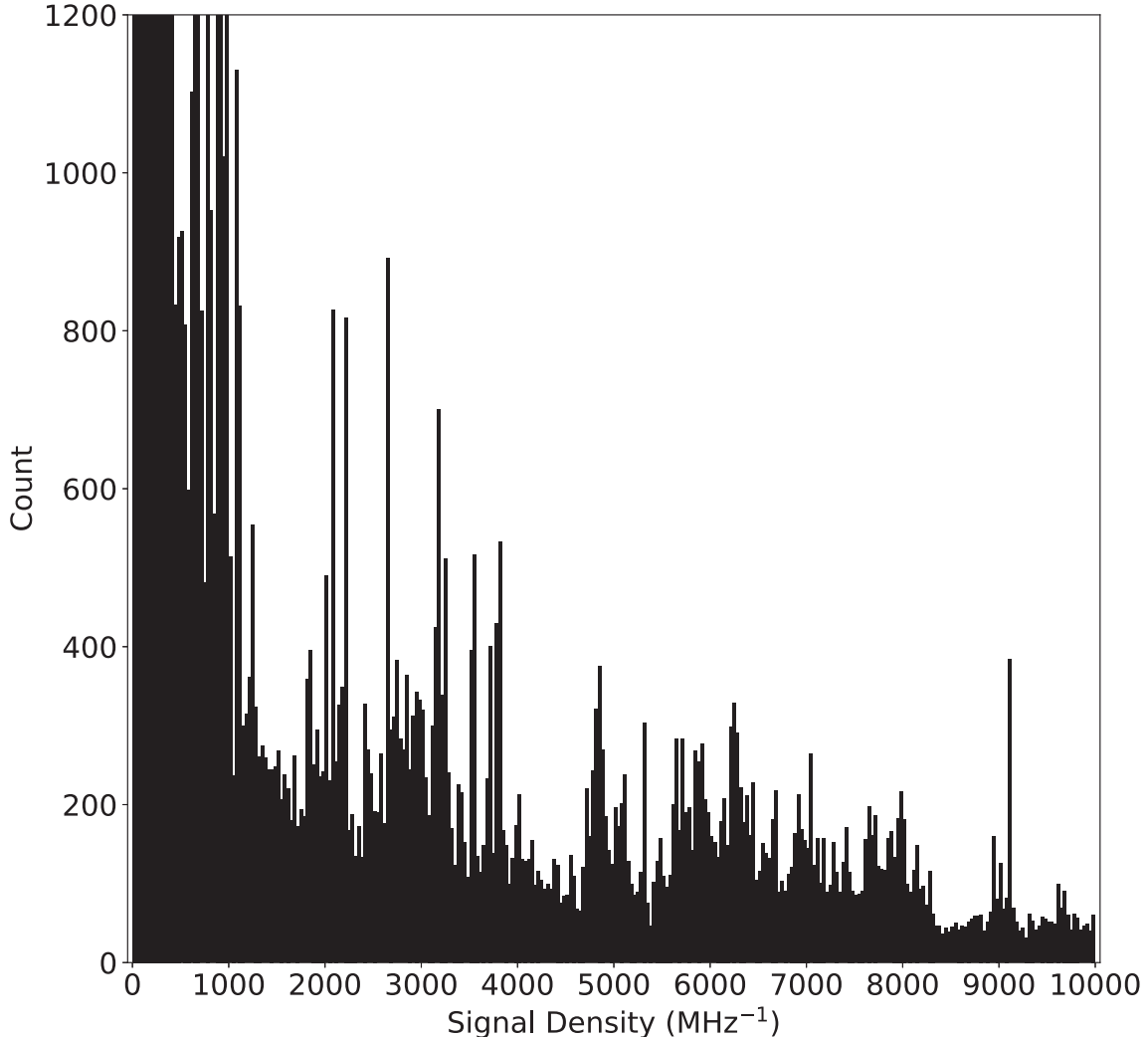

Figure 9. Signal densities of 1 kHz regions in the 1.15–1.73 GHz range. The plot is clipped at a max of 1200 signals $\mathrm{MHz^{-1}}$. Note the sharp drop-off at approximately 1000 signals $\mathrm{MHz^{-1}}$.

Appendix D
Re-analysis of 2016 Data

Margot et al. (2018) presented the results of a search for technosignatures around 14 planetary systems in the *Kepler* field conducted on 2016 April 15, 16:00–18:00 UT with the GBT. They discarded a frequency region of width $\pm \dot{f}_{max} \tau \simeq$ 3000 Hz around each detection, where $\dot{f}_{max} = 8.86\,\mathrm{Hz\,s^{-1}}$ is the maximum drift rate detectable by their search and $\tau \approx 150\,\mathrm{s}$ is the integration time of the scan. This process or a variant of it is also implemented in the radio technosignature detection pipelines of Siemion et al. (2013) and Enriquez et al. (2017). As a result, a substantial portion of the spectrum remains unexamined in these searches, and the DFM associated with these searches is overestimated (Section 5).

To remedy this situation, we have re-analyzed the data obtained by Margot et al. (2018) using the candidate detection procedure described in this work. We have found a total of 4,228,085 signals as compared with 1,046,144 detected previously. Our Doppler and direction-of-origin rejection algorithms (Section 3.4, Appendix B) automatically labeled more than 99% of the detected signals as anthropogenic RFI. After removing all of the remaining signals found within operating bands of the interferers described by Margot et al. (2018, their Tables 2 and 3), we were left with 18 technosignature candidates. Seven of these had been identified by Margot et al. (2018) and attributed to anthropogenic RFI. The properties of the remaining 11 technosignature candidates are given in Table 4. These signals were further scrutinized and categorized according to the procedure described in Section 4. All were found to be attributable to anthropogenic RFI.

We note that 3 of the 19 final technosignature candidates described by Margot et al. (2018) were found to be part of a broadband RFI signal and were removed via Equation (1). The remaining nine signals were correctly labeled as RFI by our Doppler and direction-of-origin rejection algorithms or frequency-based filters. This enhancement in classification performance was only possible because of the improvements to the candidate signal detection algorithms presented in this work and because the raw data were preserved for re-analysis.

Table 4
Characteristics of the Top 11 Candidates from the 2016 Search

Source	Epoch (MJD)	Frequency (Hz)	Drift Rate (Hz s^{-1})	S/N	Category
Kepler-22	57493.70935185	1456613901.257515	0.52042	14.3	a
Kepler-296	57493.71608796	1454808911.681175	0.05204	10.3	a
		1454877239.465714	0.01735	14.0	a
		1457440766.692162	0.31225	12.9	a
		1457443791.627884	0.31225	11.9	b
Kepler-399	57493.68870370	1414058676.362038	0.01735	764.4	b
Kepler-407	57493.71819444	1620381447.672844	−6.40113	14.5	d
Kepler-440	57493.74275463	1457486632.466316	0.41633	21.4	b
		1457486835.122108	0.39899	23.6	b
Kepler-442	57493.74497685	1308946093.916893	−0.01735	11.3	e
		1376548311.114311	−0.24286	540.9	b

Note. Properties are listed for the first scan of the source only. For a description of the category column, see Section 4.

ORCID iDs

Pavlo Pinchuk ● https://orcid.org/0000-0003-4736-4728
Jean-Luc Margot ● https://orcid.org/0000-0001-9798-1797
Adam H. Greenberg ● https://orcid.org/0000-0001-8834-9423
Ryan S. Lynch ● https://orcid.org/0000-0001-5229-7430

References

Borucki, W. J. 2016, RPPh, 79, 036901
Cohen, R. J., Downs, G., Emerson, R., et al. 1987, MNRAS, 225, 491
Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, Natur, 544, 333
Drake, F. 1984, in SETI Science Working Group Report, NASA Technical Paper No. TP-2244, ed. F. Drake, J. H. Wolfe, & C. L. Seeger (Greenbelt, MD: NASA), 67
DuPlain, R., Ransom, S., Demorest, P., et al. 2008, Proc. SPIE, 7019, 70191D
Enriquez, J. E., Siemion, A., Foster, G., et al. 2017, ApJ, 849, 104
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Natur, 542, 456
Gray, R. H., & Mooley, K. 2017, AJ, 153, 110
Harp, G. R., Richards, J., Tarter, J. C., et al. 2016, AJ, 152, 181
Kane, S. R., Hill, M. L., Kasting, J. F., et al. 2016, ApJ, 830, 1
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, AJ, 765, 131
Margot, J.-L., Greenberg, A. H., Pinchuk, P., et al. 2018, AJ, 155, 209
Savitzky, A., & Golay, M. J. E. 1964, AnaCh, 36, 1627
Siemion, A. P. V., Demorest, P., Korpela, E., et al. 2013, ApJ, 767, 94
Tarter, J. 2001, ARA&A, 39, 511
Tarter, J., Agrawal, A., Ackermann, R., et al. 2010, Proc. SPIE, 7819, 781902
Taylor, J. H. 1974, A&AS, 15, 367